Eur J Nucl Med Mol Imaging (2008) 35:1602–1607
DOI 10.1007/s00259-008-0807-0

ORIGINAL ARTICLE

A decision support system improves the interpretation of myocardial perfusion imaging

K. Tägil · M. Bondouy · J. P. Chaborel · W. Djaballah · P. R. Franken · S. Grandpierre · B. Hesse · M. Lomsky · P. Y. Marie · T. Poisson · L. Edenbrandt

Received: 27 November 2007 / Accepted: 5 April 2008 / Published online: 20 May 2008
© Springer-Verlag 2008

Abstract
Purpose The aim of this study was to investigate the influence of a computer-based decision support system (DSS) on performance and inter-observer variability of interpretations regarding ischaemia and infarction in myocardial perfusion scintigraphy (MPS).
Methods Seven physicians independently interpreted 97 MPS studies, first without and then with the advice of a DSS. Four physicians had long experience and three had limited experience in the interpretation of MPS. Each study was interpreted regarding myocardial ischaemia and infarction in five myocardial regions. The patients had undergone a gated MPS using a 2-day stress/gated rest ^{99m}Tc sestamibi protocol. The gold standard used was the interpretations made by one experienced nuclear medicine specialist on the basis of all available clinical and image information.
Results The sensitivity for ischaemia of the seven readers increased from 81% without the DSS to 86% with the DSS (p=0.01). The increase in sensitivity was higher for the three inexperienced physicians (9%) than for the four experienced physicians (2%). There was no significant change in specificity between the interpretations. The interpretations of ischaemia made with the advice of the DSS showed less inter-observer variability than those made without advice.
Conclusion This study shows that a DSS can improve performance and reduces the inter-observer variability of interpretations in myocardial perfusion imaging. Both experienced and, especially, inexperienced physicians can improve their interpretation with the advice from such a system.

Keywords Myocardial perfusion imaging · Interpretation · Computer aided diagnosis · Neural networks · Coronary artery disease

K. Tägil (✉) · L. Edenbrandt
Department of Clinical Sciences, Malmö,
Lund University Research Program in Medical Informatics,
Malmö University Hospital,
205 02 Malmö, Sweden
e-mail: kristina.tagil@med.lu.se

M. Bondouy · J. P. Chaborel · P. R. Franken
Department of Nuclear Medicine, CHU—Nice,
Nice, France

W. Djaballah · S. Grandpierre · P. Y. Marie · T. Poisson
Department of Nuclear Medicine, CHU—Nancy,
Nancy, France

B. Hesse
Department of Clinical Physiology and Nuclear Medicine,
Rigshospitalet,
Copenhagen, Denmark

M. Lomsky · L. Edenbrandt
Department of Clinical Physiology,
Sahlgrenska University Hospital,
Gothenburg, Sweden

Introduction

Computer-based decision support systems (DSS) capable of providing physicians with patient-specific advice to aid in clinical decision making could be a useful tool to improve the quality of image reporting. DSS that provides these advices automatically as part of the physician's workflow at the time and location of decision making have proved to be the most successful [1].

Several DSS for the interpretation of myocardial perfusion scintigraphy (MPS) have been presented. An expert system for the computer-assisted diagnosis of MPS has been developed since the 1980s by Garcia et al., and a study of the diagnostic performance of their system was presented in 2001 [2]. They presented a sensitivity and specificity of 83% and 73%, respectively, for overall detection of coronary artery disease, using an expert as the gold standard. This type of rule-based expert systems has achieved limited success in the medical field, and the results of Garcia again show that the interpretation of diagnostic images is a pattern recognition task, the result of which cannot easily be encapsulated in a set of criteria. Case-based reasoning and artificial neural networks are techniques in the field of artificial intelligence that have proved to be of value in pattern recognition tasks, and several applications have been presented for the interpretation of MPS [3–8]. We have recently developed an artificial neural network based DSS for the interpretation of gated single-photon emission computed tomography (SPECT) MPS [9]. Like physicians, the neural networks rely on their accumulated experience rather than on simple rules of how to interpret diagnostic images. The neural networks were trained using a database of MPS studies from the Royal Brompton Hospital in London. Features describing the function and perfusion of the left ventricle, at stress and at rest, are used as inputs to the neural networks. The performance of the DSS has been evaluated in a different hospital from the one in which it was trained, and the results were encouraging [9]. However, our DSS is not intended to replace the physician, and so the performance of the DSS per se is not the most important factor. It is more important to know whether the physician will benefit from the advice given by the DSS.

The aim of this study was to investigate the influence of a computer-based DSS on performance and inter-observer variability of interpretations regarding ischaemia and infarction in myocardial perfusion imaging.

Materials and methods

Patients

All patients who had undergone electrocardiogram (ECG)-gated MPS between May 4 and September 14, 2005, at Sahlgrenska University Hospital in Gothenburg were studied retrospectively. These patients were part of a larger material previously used to validate the image processing algorithms of the DSS [10]. Images with sub-optimal technical quality were excluded to focus the observers on the interpretation of different perfusion patterns. The DSS was not able to automatically differentiate the left ventricle from extra cardiac activity in 16 patients, and manual steps would have been required. These cases were excluded as were 23 patients with poor technical image quality, which may have prevented some of the observers to interpret the images. No patients were excluded due to the appearance of the perfusion pattern or the interpretation presented by the DSS. One hundred patients were selected for the study, but in the process of preparing the CDs for the observers, three cases were not properly transferred making the final study material comprising 97 patients. The mean age of the patients was 67±6 years, and 55 were men and 42 women. All patients were referred to a MPS for the diagnosis or management of coronary artery disease. Twelve of the patients were smokers, prior myocardial infarction was found in 15 patients, chest pain in 41, family history of ischaemic heart disease in 32, diabetes in 12, peripheral vascular disease in four, hypertension in 46, hyperlipidaemia in 48, prior percutanerous transluminal coronary angioplasty in 18 and prior coronary artery bypass graft surgery in nine patients. None of the patients in the study group were included in the data set used for training the DSS. The study was approved by the Research Ethics Committee at Gothenburg University.

Radionuclide Imaging

The gated SPECT studies were performed using a 2-day non-gated stress/gated rest ^{99m}Tc sestamibi protocol. Patients were stressed using either maximal exercise (50 of 97) or a pharmacological test with adenosine (47 of 97). The exercise was continued for at least 2 min and the adenosine infusion for at least 2.5 min, after the injection of the tracer. Stress and rest image acquisition began about 60 min after the injection of approximately 600 MBq ^{99m}Tc sestamibi. Images were acquired with one of two different dual-head SPECT cameras (Infinia or Hawkeye, General Electric, USA) with the detectors in an L-shape position equipped with low-energy high-resolution collimators.

Acquisition was performed in the step-and-shoot mode over 90° with 30 projections per detector and 40 s per projection with the patient in the supine position, using circular rotation and a 64×64 matrix, zoom factor 1.28 and pixel size 6.9 mm. In patients weighing more than 90 kg, the acquisition time per projection was increased to 55 s. During the rest acquisition, the patient was monitored with a three-lead ECG. The acceptance window was opened to ±20% of the pre-defined R–R interval. Other beats were rejected for ejection fraction and volume calculation. Each R–R interval was divided into eight equal time intervals. Gated SPECT acquisition was performed at the same time as the ungated routine SPECT acquisition. An automatic motion correction programme was applied where patient motion during acquisition was apparent. Tomographic reconstruction of non-gated data was performed using filtered back-projection with a Butterworth filter with a critical frequency of 0.52 cycles per centimetre and order 5.

The reconstruction of gated data was achieved using filtered back-projection with a Butterworth filter with a critical frequency of 0.40 cycles per centimetre and order 10. No attenuation or scatter correction was used.

Gold standard

The clinical interpretations of each MPS by one senior nuclear medicine specialist with more than 25 years of experience in nuclear cardiology was used as the gold standard regarding presence of ischaemia and infarction in five left ventricular regions (anterior, lateral, septal, inferior and apical region). These interpretations were based on the combined information from the images, including quantitative data from the Emory Cardiac Toolbox analytical software package [11, 12], as well as clinical information.

Decision Support System

The EXINI heart software package (version 2.2; Exini Diagnostics AB, Lund, Sweden) was used to display the MPS images and to present the advice of the DSS to the observers. The DSS used for automated interpretation of the gated MPS images was based on automated image processing algorithms, artificial neural networks and a database of classified MPS images from the Royal Brompton Hospital in London [9,10].

Observers

Seven physicians participated in the study. All were specialists in nuclear medicine, but with varying experience in nuclear cardiology; four were regarded as experienced observers (3, 10, 17 and 20 years of practise) and three as inexperienced (less than 1 year of practise). The gold standard reader was not one of the seven observers. Each observer interpreted the 97 MPS cases independently, without knowledge of the interpretations made by the other observers.

Each MPS was classified regarding absence and presence of myocardial ischaemia and myocardial infarction in each of five left ventricular regions anterior, septal, inferior, lateral and apical. First, the observer made an interpretation without advice from the DSS regarding myocardial ischaemia and/or infarction in each of the five myocardial regions. These interpretations were registered in the Exini software. As the next step, the DSS presented the diagnostic advice for each of the ten classifications. If there were disagreements between the observer and the DSS, the observer had the option to keep his/her first interpretation or accepting the interpretation proposed by the DSS. The advice and the final interpretation were also registered in the Exini software.

Based on the regional classifications, an interpretation regarding left ventricular infarction and/or ischaemia was defined. If infarction or ischaemia was found in at least one of the regions, left ventricular infarction or ischaemia was present, and if infarction and ischaemia were absent in all five regions, left ventricular infarction and ischaemia were absent.

Statistical methods

The influence of the DSS on performance was analysed in the following way. If an observer made a false interpretation (compared to the gold standard) in his/her initial interpretation without DSS and a true interpretation in the final interpretation after the advice of the DSS, that case showed ‘improved performance’ and vice versa. The number of cases showing increased and decreased performance was counted. The significance of changes was evaluated using a McNemar type of statistics.

The influence of the DSS on inter-observer variability was analysed in the following way. For each of the 97 patients, the highest number of observers that made the same interpretation was calculated, both with and without DSS. Variability was defined to decrease if this number was higher with DSS compared to without DSS and vice versa. If only one of seven observers without DSS classified a patient as having ischaemia and he changed his classification to ‘no ischaemia’ with the advice of the DSS, so that all seven observers agreed, then the DSS decreased the variability of for that patient. The numbers of decreased and increased variability were compared. The significance of an imbalance in these numbers was evaluated using a McNemar type of statistics with $p<0.05$ being considered as significant.

Results

The performances of the inexperienced and experienced observers in their interpretations ‘without’ and ‘with’ DSS regarding myocardial ischaemia and infarction are presented in Tables 1 and 2. With advice of the DSS, the seven observers increased their sensitivity for left ventricular ischaemia with 5% units (81% to 86%; $p<0.01$). The specificity was unchanged. The increase in sensitivity was greater for the three inexperienced observers (9% units; from 78% to 87%; $p=0.02$) than for the four experienced observers (2% units; from 84% to 86%; n.s.). On a regional level, a statistically significant increase in sensitivity with DSS was found in the inferior (10%; $p=0.004$) and anterior (7%; $p=0.03$) regions. There were no significant changes in specificity for ischaemia between the interpretations without and with DSS. The DSS showed a higher sensitivity for left ventricular ischaemia than the seven observers (100% vs. 81%), while the observers showed a higher specificity than the DSS (71% vs. 66%).

Table 1 Results of the seven observers in their interpretations 'without DSS' and 'with DSS' of 97 myocardial perfusion scintigrams regarding ischaemia

Ischaemia	Left ventricular, *n*=23		Anterior, *n*=13		Septal, *n*=5		Inferior, *n*=13		Lateral, *n*=10		Apical, *n*=14	
	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp
Interpretation without DSS												
Inexperienced	78	67	74	88	80	93	67	81	57	89	60	91
Experienced	84	75	79	91	85	96	75	86	78	93	82	86
All	81	71	77	89	83	94	71	84	69	91	72	88
Changes of interpretation with DSS												
Inexperienced	9	−1	13	1	7	1	8	0	10	0	2	−3
Experienced	2	−2	2	−1	5	0	12	−1	0	0	2	2
All	5	0	7	0	6	0	10	−1	4	0	2	0

Values shown as percentages

Se Sensitivity, *Sp* specificity, *DSS* computer-based decision support system, *n* number of abnormal cases, *Inexperienced* <1 year of practise, *Experienced* 3–20 years of practise

The observers' sensitivity for infarction was 96% for all observers without an advice from the DSS and 97% with the advice from the DSS (n.s.). The specificity was unchanged. The DSS showed higher specificity for left ventricular infarction than the seven observers (82% vs. 70%), while the observers showed a higher sensitivity than the DSS (96% vs. 93%).

In 46 patients, the DSS analysis gave the diagnosis of 'no ischaemia.' This was in agreement with the gold standard in all cases, giving a negative predictive value of 100%. DSS analysis also ruled out infarction in 69 patients, one of them having an infarction, which gave a negative predictive value for infarction of 98.6%.

Regarding left ventricular ischaemia, the two most experienced observers agreed with each other in 74 of the 97 patients and with the gold standard in 84 and 71 of the patients, respectively; that is, the inter-observer variations between the 'gold-standard observer' and the two other experienced physicians were similar and fairly low.

The inter-observer variability decreased for inferior infarction with the advice of the DSS (decreased variability in nine patients and increased in one; $p=0.01$). The lowest number of patients in which all seven observers agreed was found in the inferior region (44 for infarction and 43 for ischaemia). In the other regions, the corresponding figures were higher for infarction (73–83) than for ischaemia (51–64; Tables 3 and 4).

Discussion

Main findings

The results of this study show that the physicians benefited from the advice of a DSS, in terms of both improved performance and decreased inter-observer variability. The provision of advice from the DSS resulted in a significant improvement in sensitivity for ischaemia in the anterior and

Table 2 Results of the seven observers in their interpretations 'without DSS' and 'with DSS' of 97 myocardial perfusion scintigrams regarding infarction

Infarction	Left ventricular, *n*=13		Anterior, *n*=4		Septal, *n*=4		Inferior, *n*=11		Lateral, *n*=5		Apical, *n*=7	
	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp	Se	Sp
Interpretation without DSS												
Inexperienced	98	57	100	90	83	95	97	67	87	93	62	88
Experienced	95	79	88	97	75	97	91	83	80	97	64	93
All	96	70	93	94	79	96	94	76	83	95	63	91
Changes of interpretation with DSS												
Inexperienced	2	0	0	0	8	−1	3	1	0	−1	5	0
Experienced	0	0	13	−1	0	0	0	0	−10	0	0	1
All	1	0	7	0	4	0	1	0	−6	0	2	0

Values shown as percentages

Se Sensitivity, *Sp* specificity, *DSS* computer-based decision support system, *n* number of abnormal cases, *Inexperienced* <1 year of practise, *Experienced* 3–20 years of practise

Table 3 Influence of DSS on inter-observer variability for seven observers interpreting 97 myocardial perfusion scintigrams regarding ischaemia

Ischaemia	Left ventricular	Anterior	Septal	Inferior	Lateral	Apical
Agreement	33	54	64	43	54	51
Disagreement						
Variability ↓	10	7	7	16	8	9
Variability ↑	7	8	4	7	4	3
Variability →	47	28	22	31	31	34
Total	97	97	97	97	97	97

Agreement All seven physicians agreed, *Disagreement* at least one of the seven physician disagreed, ↓ decreased, ↑ increased, → unchanged

inferior region but not for infarction. One explanation for this could be that the number of regions with infarction was much lower than the number with ischaemia (31 vs. 55). Another explanation could be that the interpretation of ischaemia is more difficult than that of infarction. The presence or absence of ischaemia is generally the clinically most important part of the interpretation, and the benefit of a DSS for the interpretation of ischaemia would therefore be of special value.

In a previous study from our group, we showed decreased inter-observer and intra-observer variability as well as improved performance [13]. This is in accordance with the results of the present study. The design of this study is closer to a clinical situation, in that the observers analysed slice images (rest, stress and gated rest), polar plots and three-dimensional images as well as the results of the quantification of left ventricular ejection fraction and volumes. In the previous study, the observers only had polar plots to interpret. Another difference is that the observers in the previous study interpreted the cases in terms of the presence or absence of coronary artery disease, whereas the present study was focused on the type of interpretation that physicians make in clinical practise; that is, the presence or absence of ischaemia and infarction. Additionally, in the previous study, the gold standard classification of coronary artery disease was based on an independent method, coronary angiography. Coronary angiography has often been used as a gold standard in MPS studies, but its disadvantage is that it can only visualize coronary artery stenosis rather than myocardial perfusion. Accepted practise is therefore moving towards the use of expert interpretations regarding ischaemia and infarction. The gold standard in the present study was the classifications of a very experienced physician who made an assessment based not only on the images but also on patient history, risk factors, results from earlier examinations, data from the stress test and, if present, coronary angiography. A potential problem with using a single physician for the gold-standard classification could arise if this person has a different ‘style’ of interpreting MPS images compared to the other observers. We therefore tested whether the two most experienced observers agreed more with each other than with the gold standard classification; this was not the case.

In our previous study, we included three physicians trained and working at the same hospital. The inter-observer variability between the physicians of such a group is probably less than that between physicians from different hospitals. In this study, we included seven observers from two different hospitals.

Clinical implications

This study shows that a DSS based on automated image processing and artificial neural networks can support and improve the interpretation of MPS. The DSS could be seen as an experienced colleague who gives a second opinion to the reporting physician. Unlike a computer, however, there is not always an experienced colleague available at the time and place where the advice is needed. A DSS also has the additional advantage of being more consistent and presents more standardized interpretations. Lindahl et al. showed that even a highly experienced physician made different

Table 4 Influence of DSS on inter-observer variability for seven observers interpreting 97 myocardial perfusion scintigrams regarding infarction

Infarction	Left ventricular	Anterior	Septal	Inferior	Lateral	Apical
Agreement	39	73	83	44	74	73
Disagreement						
Variability ↓	8	4	4	9	4	3
Variability ↑	5	3	1	1	6	2
Variability →	45	17	9	43	13	19
Total	97	97	97	97	97	97

Agreement All seven physicians agreed, *Disagreement* at least one of the seven physician disagreed, ↓ decreased, ↑ increased, → unchanged interpretations for the same MPS case on different occasions and that the use of a DSS reduced this intra-observer variability [13].

Conclusion

This study shows that a DSS can be used to improve performance and decrease inter-observer variability regarding interpretations in myocardial perfusion imaging. After advice of a DSS, physicians will interpret MPS better compared to a gold standard of consensus reading by trained nuclear cardiologists. Even experienced physicians improved their sensitivity after advice from the DSS. In the near future, this type of DSS may become a valuable tool to improve clinical practise in diagnostic imaging.

Acknowledgements This study was supported by a scholarship from Medicon Valley Association.

Conflicts of interests statement Kristina Tägil and Lars Edenbrandt are shareholders in Exini Diagnostics AB, which owns the EXINI heart software used in the study.

References

1. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 2005;330:765.
2. Garcia EV, Cooke CD, Folks RD, Santana CA, Krawczynska EG, Braal LD, Ezquerra NF. Diagnostic performance of an expert system for the interpretation of myocardial perfusion SPECT studies. J Nucl Med 2001;42:1185–91.
3. Haddad M, Adlassnig KP, Porenta G. Feasibility analysis of a case-based reasoning system for automated detection of coronary heart disease from myocardial scintigrams. Artif Intell Med 1997;9:61–78.
4. Khorsand A, Graf S, Sochor H, Schuster E, Porenta G. Automated assessment of myocardial SPECT perfusion scintigraphy: a comparison of different approaches of case-based reasoning. Artif Intell Med 2007;40:103–13.
5. Fujita H, Katafuchi T, Uehara T, Nishimura T. Application of artificial neural network to computer-aided diagnosis of coronary artery disease in myocardial SPECT bull's-eye images. J Nucl Med 1992;33:272–6.
6. Porenta G, Dorffner G, Kundrat S, Petta P, Duit-Schedlmayer J, Sochor H. Automated interpretation of planar thallium-201-dipyridamole stress-redistribution scintigrams using artificial neural networks. J Nucl Med 1994;35:2041–7.
7. Hamilton D, Riley PJ, Mioloa UJ, Amro AA. A feed forward neural network for classification of bull's-eye myocardial perfusion images. Eur J Nucl Med 1995;22:108–15.
8. Lindahl D, Palmer J, Ohlsson M, Peterson C, Lundin A, Edenbrandt L. Automated interpretation of myocardial SPECT perfusion images using artificial neural networks. J Nucl Med. 1997;38:1870–5.
9. Gjertsson P, Lomsky M, Richter J, Ohlsson M, Tout D, van Aswegen A, et al. The added value of ECG-gating for the diagnosis of myocardial infarction using myocardial perfusion scintigraphy and artificial neural networks. Clin Physiol Funct Imaging 2006;26:301.
10. Lomsky M, Richter J, Johansson L, Hoilund-Carlsen PF, Edenbrandt L. Validation of a new automated method for analysis of gated-SPECT images. Clin Physiol Funct Imaging 2006;26:139.
11. Van Train KF, Areeda J, Garcia EV, Cooke CD, Maddahi J, Kiat H, et al. Quantitative same-day rest-stress technetium-99m-sestamibi SPECT: definition and validation of stress normal limits and criteria for abnormality. J Nucl Med 1993;34:1494.
12. Van Train KF, Garcia EV, Maddahi J, Areeda J, Cooke CD, Kiat H, et al. Multicenter trial validation for quantitative analysis of same-day rest-stress technetium-99m-sestamibi myocardial tomograms. J Nucl Med 1994;35:609.
13. Lindahl D, Lanke J, Lundin A, Palmer J, Edenbrandt L. Improved classifications of myocardial bull's-eye scintigrams with computer-based decision support system. J Nucl Med 1999;40(1):96–101.